ROSALES, Angel Abegail B.

BSIT 3-5

What is Data Science?

Data science is the study of data to extract meaningful insights for business. It is a multidisciplinary approach that combines principles and practices from the fields of mathematics, statistics, artificial intelligence, and computer engineering to analyze large amounts of data.

Differentiate between Business Intelligence (BI) and Data Science

Business Intelligence (BI) - focuses on analyzing historical and current data to provide actionable insights for business operations. It is primarily descriptive and diagnostic, answering "what happened" and "why it happened" by organizing and visualizing data.

Data Science - goes beyond just analyzing past data; it also uses advanced statistical techniques, machine learning, and algorithms to predict future outcomes and generate new patterns or trends. It answers questions like "what will happen next" and "how can we achieve a specific outcome" through predictive and prescriptive analysis.

The Role of Data Visualization in Data Science

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets.

Data visualization is important for almost every professional discipline. Teachers use it to display student test results, executives to share information with stakeholders, etc.

Visualization tools are a natural fit to provide useful information.

Give me at least three Fundamentals Concept of Data Science and explain each.

Algorithms - A specific set of rules or processes used in a calculation to solve problems or perform a task. For example, a recipe - a set of rules to follow to get a specific outcome. It is used for data analysis and data modeling, among other functions.

Machine Learning - AI programming of a system to automatically perform a specific task, self-learn from the data collected, perform pattern recognition and make decisions, with little to no human intervention. It is use to build predictive models.

Programming - Programming languages are used to develop and build data models, clean and organize data and visualize the data. Popular languages are Python, R and SQL.